Title:	Prototyping FLAG Monitoring in Splunk
Author(s):	Tucci, Emily Patrina
Intended for:	Report
Issued:	2020-08-11

Disclaimer:
Los Alamos National Laboratory, an affirmative action/equal opportunity employer, is operated by Triad National Security, LLC for the National Nuclear Security Administration of U.S. Department of Energy under contract 89233218CNA000001. By approving this article, the publisher recognizes that the U.S. Government retains nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or to allow others to do so, for U.S. Government purposes. Los Alamos National Laboratory requests that the publisher identify this article as work performed under the auspices of the U.S. Department of Energy. Los Alamos National Laboratory strongly supports academic freedom and a researcher's right to publish; as an institution, however, the Laboratory does not endorse the viewpoint of a publication or guarantee its technical correctness.

Prototyping FLAG Monitoring in Splunk

(My Summer Internship at LANL)

Emily Tucci

Aug 5th 2020

Los Alamos
NATIONAL LABORATORY
EST. 1943

NNSA
National Nuclear Security Administration

Managed by Triad National Security, LLC for the U.S. Department of Energy's NNSA

Outline

Background

Motivation

Toolbox

Workflow

Results: Tracking & Reporting in Splunk

Wrap Up/Future Work/Questions

Who am I?

Rising Senior @ Oberlin College (40 min SW Cleveland)

- Majors: CS, Math, Hispanic Studies

Work: CS-TA, Tutor Children, Gym

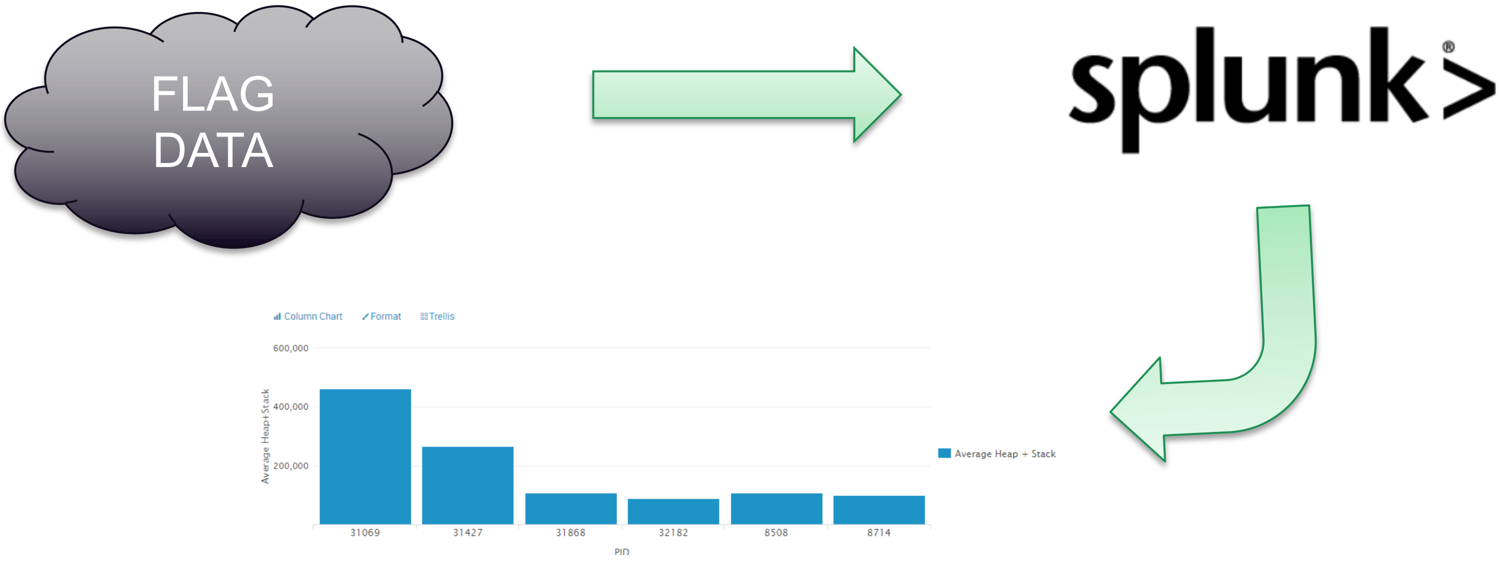
MEMORY BATTLES OF THE SPANISH CIVIL WAR
HISTORY, FICTION, PHOTOGRAPHY
SEBASTIAAN FABER

Favorite Hobby: (Softball/Pitcher)

Overall Goal

2 parts:

- Send new memory data to Splunk
- Utilize Splunk commands to analyze + interpret FLAG data



FLAG DATA

What did I add to my Toolbox?

Technical:

- Parallel computing
- Building + Running FLAG
- How to work Splunk (w/ FLAG specific data)

Professional:

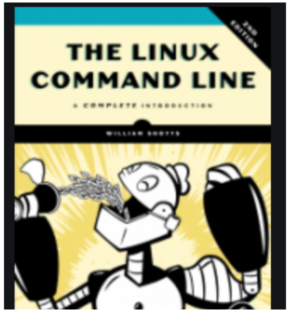
- Wider view of work done at the lab
- Better understanding of how to compare industry and lab/academia
- How grad school could meaningfully fit into my goals

Workflow

Week 1: Linux Refresh + Parallel Computing

Week 2-3: What is FLAG + Intro to HPC

Week 4-5: Splunk Training/Navigate X-Splunk



2ND EDITION

THE LINUX COMMAND LINE

A COMPLETE INTRODUCTION

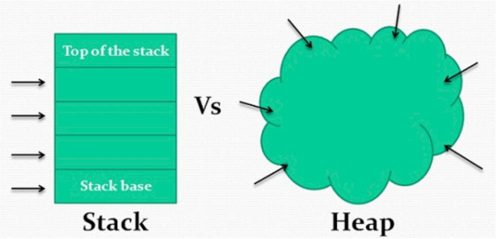
WILLIAM SHOTTS

Week 6: Existing timers in FLAG + develop queries

Week 7-8: Analyze MemInfo Reporting

Week 9-10: Add ***print_run_info()*** calls + analyze in Splunk

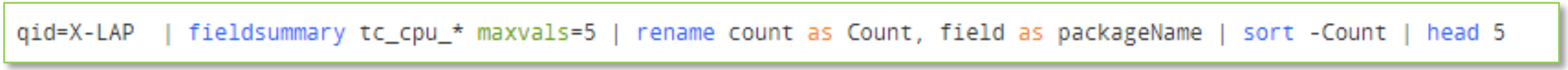
splunk>®



Internal Timers + Splunk

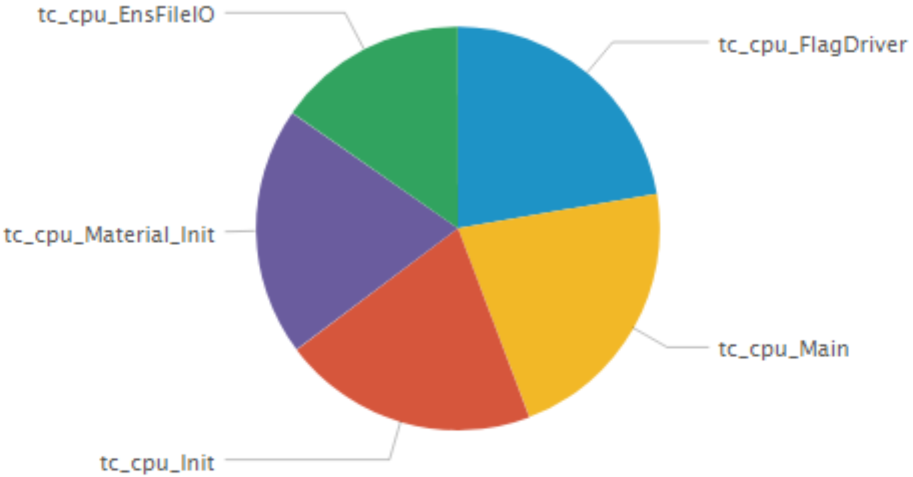
Before Diving Into MemInfo:

- Used pre-existing fields: tc_cpu_<packagename>
- Tracking & Reporting (Top 5)



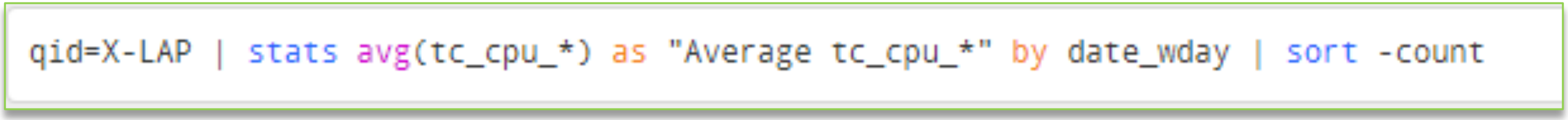
```
qid=X-LAP  | fieldsummary tc_cpu_* maxvals=5 | rename count as Count, field as packageName | sort -Count | head 5
```

Count/Occurrence

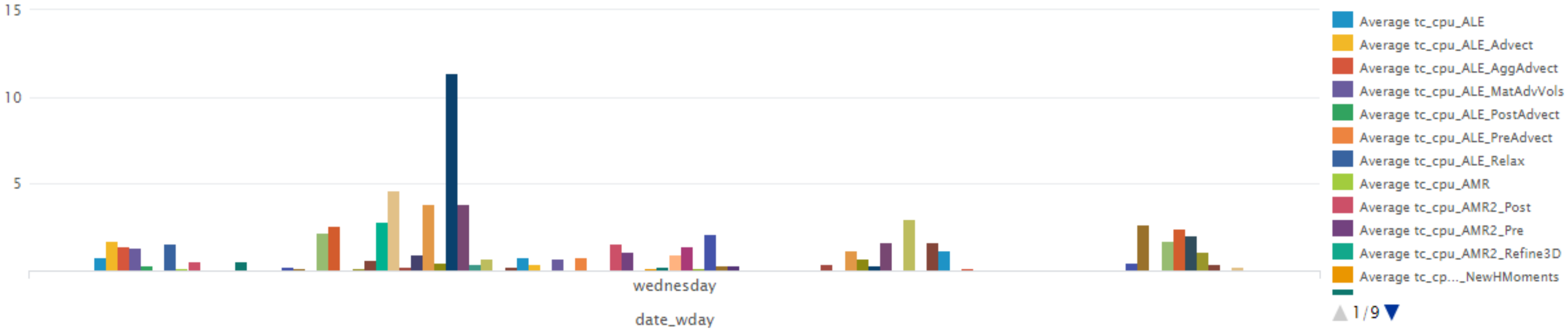


Internal Timers + Splunk (cont)

Average Time Data (seconds)



```
qid=X-LAP | stats avg(tc_cpu_*) as "Average tc_cpu_*" by date_wday | sort -count
```



Memory: Data Examined and Why

Where did we start?

- **MemInfo.cc** → lots of memory data printed to log files
  - reporting turned on via input file
- **statm** command

What value did we choose?

- heap+stack (MB)
  - variation
  - memory data is important to FLAG

How did we investigate?

- Variety of input files + settings
- Lagrange Hydro

FLAG Code Details

What did we know?

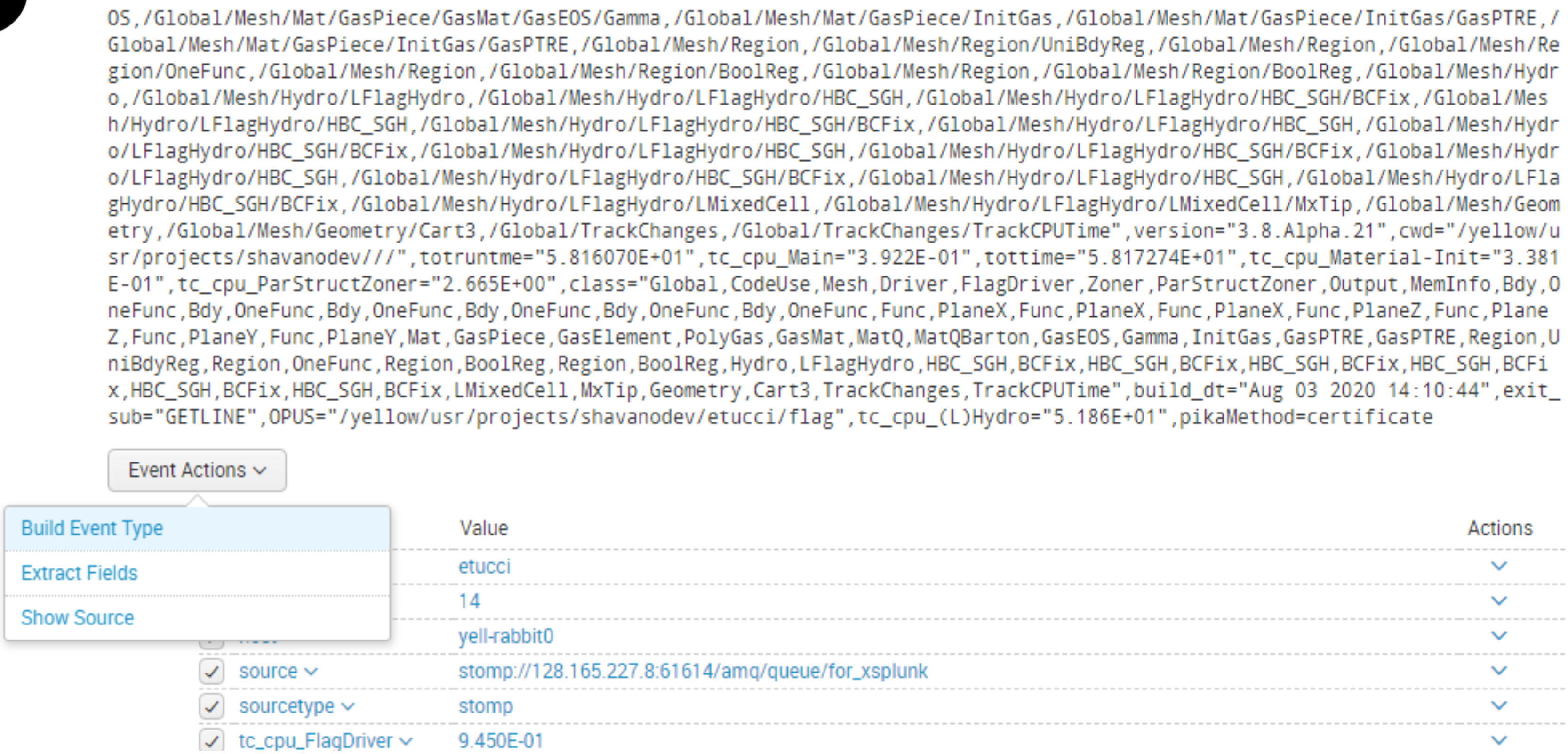
- heap+stack value in log files
- ***print_run_info(key, length, val, length)*** → sends to Splunk
- Splunk organizes by events (runs)

What do we want?

- **field 1**: value to distinguish runs
  - ***print_run_info("pid", 3, pid, strlen(pid))***
- **field 2**: value for memory data
  - ***print_run_info("heapstackpe0", 12, vdata, strlen(vdata))***

Splunk Field Extraction

1

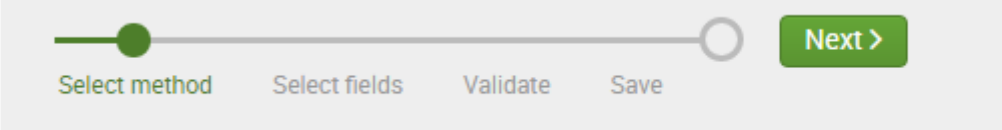


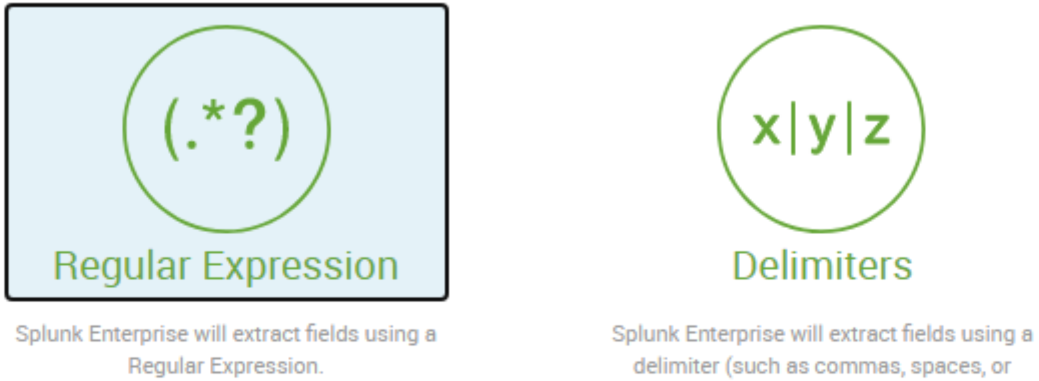
```
OS,/Global/Mesh/Mat/GasPiece/GasMat/GasEOS/Gamma,/Global/Mesh/Mat/GasPiece/InitGas,/Global/Mesh/Mat/GasPiece/InitGas/GasPTRE,/
Global/Mesh/Mat/GasPiece/InitGas/GasPTRE,/Global/Mesh/Region,/Global/Mesh/Region/UniBdyReg,/Global/Mesh/Region,/Global/Mesh/Re
gion/OneFunc,/Global/Mesh/Region,/Global/Mesh/Region/BoolReg,/Global/Mesh/Region,/Global/Mesh/Region/BoolReg,/Global/Mesh/Hydr
o,/Global/Mesh/Hydro/LFlagHydro,/Global/Mesh/Hydro/LFlagHydro/HBC_SGH,/Global/Mesh/Hydro/LFlagHydro/HBC_SGH/BCFix,/Global/Mes
h/Hydro/LFlagHydro/HBC_SGH,/Global/Mesh/Hydro/LFlagHydro/HBC_SGH/BCFix,/Global/Mesh/Hydro/LFlagHydro/HBC_SGH,/Global/Mesh/Hydr
o/LFlagHydro/HBC_SGH/BCFix,/Global/Mesh/Hydro/LFlagHydro/HBC_SGH,/Global/Mesh/Hydro/LFlagHydro/HBC_SGH/BCFix,/Global/Mesh/Hydr
o/LFlagHydro/HBC_SGH,/Global/Mesh/Hydro/LFlagHydro/HBC_SGH/BCFix,/Global/Mesh/Hydro/LFlagHydro/HBC_SGH,/Global/Mesh/Hydro/LFla
gHydro/HBC_SGH/BCFix,/Global/Mesh/Hydro/LFlagHydro/LMixedCell,/Global/Mesh/Hydro/LFlagHydro/LMixedCell/MxTip,/Global/Mesh/Geom
etry,/Global/Mesh/Geometry/Cart3,/Global/TrackChanges,/Global/TrackChanges/TrackCPUTime",version="3.8.Alpha.21",cwd="/yellow/u
sr/projects/shavanodev///",totruntme="5.816070E+01",tc_cpu_Main="3.922E-01",tottime="5.817274E+01",tc_cpu_Material-Init="3.381
E-01",tc_cpu_ParStructZoner="2.665E+00",class="Global,CodeUse,Mesh,Driver,FlagDriver,Zoner,ParStructZoner,Output,MemInfo,Bdy,O
neFunc,Bdy,OneFunc,Bdy,OneFunc,Bdy,OneFunc,Bdy,OneFunc,Bdy,OneFunc,Func,PlaneX,Func,PlaneX,Func,PlaneX,Func,PlaneZ,Func,Plane
Z,Func,PlaneY,Func,PlaneY,Mat,GasPiece,GasElement,PolyGas,GasMat,MatQ,MatQBarton,GasEOS,Gamma,InitGas,GasPTRE,GasPTRE,Region,U
niBdyReg,Region,OneFunc,Region,BoolReg,Region,BoolReg,Hydro,LFlagHydro,HBC_SGH,BCFix,HBC_SGH,BCFix,HBC_SGH,BCFix,HBC_SGH,BCFi
x,HBC_SGH,BCFix,HBC_SGH,BCFix,LMixedCell,MxTip,Geometry,Cart3,TrackChanges,TrackCPUTime",build_dt="Aug 03 2020 14:10:44",exit_
sub="GETLINE",OPUS="/yellow/usr/projects/shavanodev/etucci/flag",tc_cpu_(L)Hydro="5.186E+01",pikaMethod=certificate
```

	Value	Actions
	etucci	⌄
	14	⌄
	yell-rabbit0	⌄
☑ source ⌄	stomp://128.165.227.8:61614/amq/queue/for_xsplunk	⌄
☑ sourcetype ⌄	stomp	⌄
☑ tc_cpu_FlagDriver ⌄	9.450E-01	⌄

Splunk Field Extraction

❷





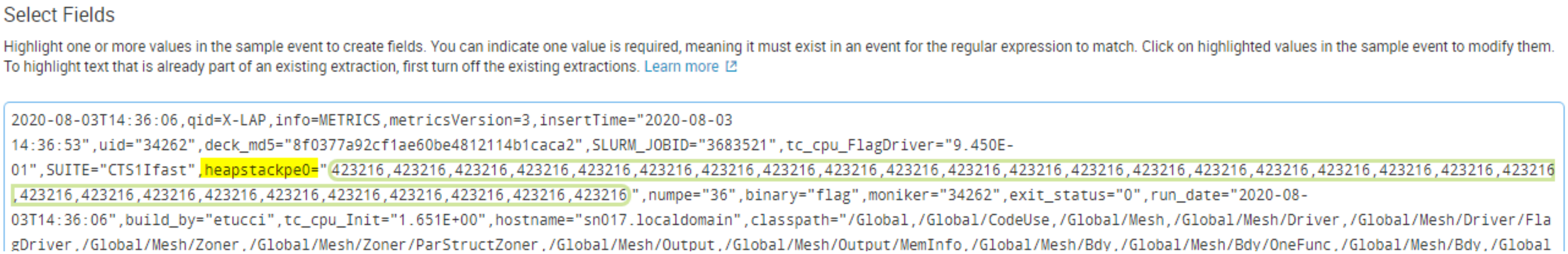
Splunk Enterprise will extract fields using a Regular Expression.

Splunk Enterprise will extract fields using a delimiter (such as commas, spaces, or

Splunk Field Extraction

KEYWORD: heapstackpe0

❸



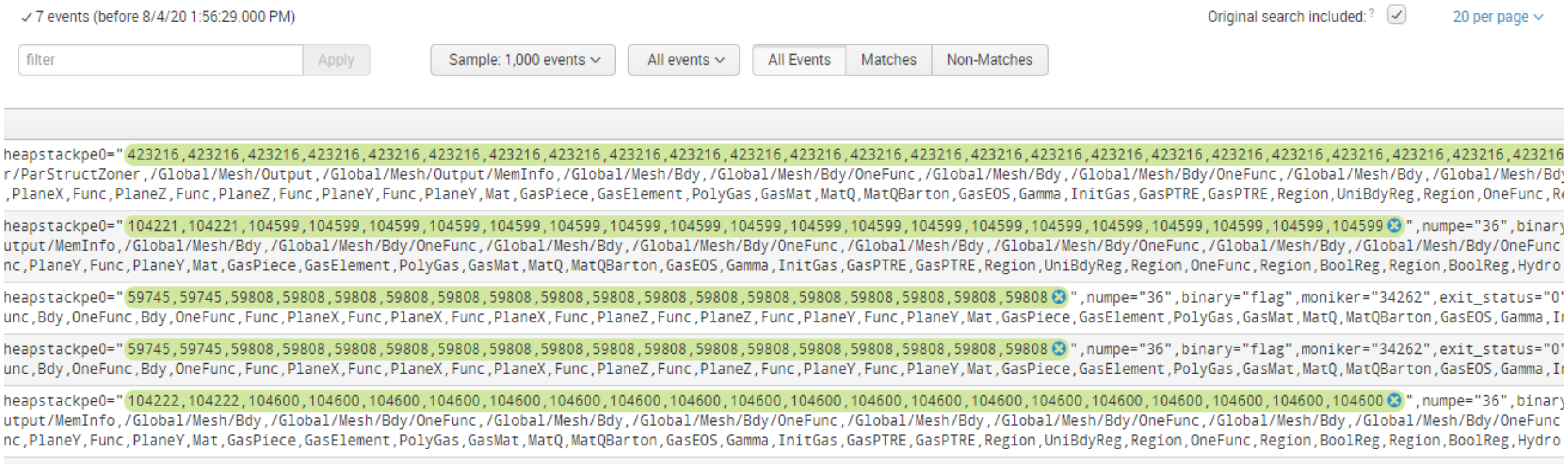
Select Fields

Highlight one or more values in the sample event to create fields. You can indicate one value is required, meaning it must exist in an event for the regular expression to match. Click on highlighted values in the sample event to modify them. To highlight text that is already part of an existing extraction, first turn off the existing extractions. Learn more

```
2020-08-03T14:36:06,qid=X-LAP,info=METRICS,metricsVersion=3,insertTime="2020-08-03
14:36:53",uid="34262",deck_md5="8f0377a92cf1ae60be4812114b1caca2",SLURM_JOBID="3683521",tc_cpu_FlagDriver="9.450E-
01",SUITE="CTS1Ifast",heapstackpe0="423216,423216,423216,423216,423216,423216,423216,423216,423216,423216,423216,423216,423216,423216,423216,423216,423216,423216,423216,423216
,423216,423216,423216,423216,423216,423216,423216,423216,423216,423216",numpe="36",binary="flag",moniker="34262",exit_status="0",run_date="2020-08-
03T14:36:06",build_by="etucci",tc_cpu_Init="1.651E+00",hostname="sn017.localdomain",classpath="/Global,/Global/CodeUse,/Global/Mesh,/Global/Mesh/Driver,/Global/Mesh/Driver/Fla
gDriver,/Global/Mesh/Zoner,/Global/Mesh/Zoner/ParStructZoner,/Global/Mesh/Output,/Global/Mesh/Output/MemInfo,/Global/Mesh/Bdy,/Global/Mesh/Bdy/OneFunc,/Global/Mesh/Bdy,/Global
```

Extraction Results

4

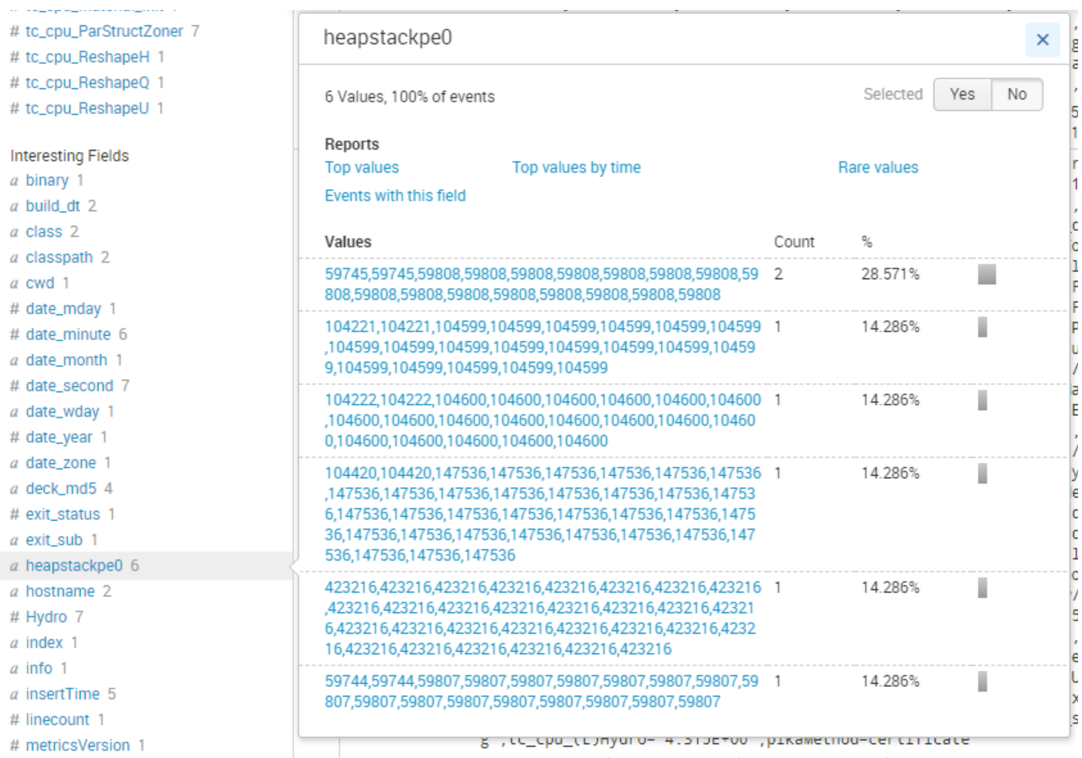


✓ 7 events (before 8/4/20 1:56:29.000 PM)

```
heapstackpe0="423216,423216,423216,423216,423216,423216,423216,423216,423216,423216,423216,423216,423216,423216,423216,423216,423216,423216,423216,423216,423216,423216,423216,423216
r/ParStructZoner,/Global/Mesh/Output,/Global/Mesh/Output/MemInfo,/Global/Mesh/Bdy,/Global/Mesh/Bdy/OneFunc,/Global/Mesh/Bdy,/Global/Mesh/Bdy/OneFunc,/Global/Mesh/Bdy,/Global/Mesh/Bdy
,PlaneX,Func,PlaneZ,Func,PlaneZ,Func,PlaneY,Func,PlaneY,Mat,GasPiece,GasElement,PolyGas,GasMat,MatQ,MatQBarton,GasEOS,Gamma,InitGas,GasPTRE,GasPTRE,Region,UniBdyReg,Region,OneFunc,Re

heapstackpe0="104221,104221,104599,104599,104599,104599,104599,104599,104599,104599,104599,104599,104599,104599,104599,104599,104599,104599,104599,104599,104599",numpe="36",binary
utput/MemInfo,/Global/Mesh/Bdy,/Global/Mesh/Bdy/OneFunc,/Global/Mesh/Bdy,/Global/Mesh/Bdy/OneFunc,/Global/Mesh/Bdy,/Global/Mesh/Bdy/OneFunc,/Global/Mesh/Bdy,/Global/Mesh/Bdy/OneFunc,
nc,PlaneY,Func,PlaneY,Mat,GasPiece,GasElement,PolyGas,GasMat,MatQ,MatQBarton,GasEOS,Gamma,InitGas,GasPTRE,GasPTRE,Region,UniBdyReg,Region,OneFunc,Region,BoolReg,Region,BoolReg,Hydro,

heapstackpe0="59745,59745,59808,59808,59808,59808,59808,59808,59808,59808,59808,59808,59808,59808,59808,59808,59808,59808",numpe="36",binary="flag",moniker="34262",exit_status="0"
unc,Bdy,OneFunc,Bdy,OneFunc,Func,PlaneX,Func,PlaneX,Func,PlaneX,Func,PlaneZ,Func,PlaneZ,Func,PlaneY,Func,PlaneY,Mat,GasPiece,GasElement,PolyGas,GasMat,MatQ,MatQBarton,GasEOS,Gamma,Ir

heapstackpe0="59745,59745,59808,59808,59808,59808,59808,59808,59808,59808,59808,59808,59808,59808,59808,59808,59808,59808",numpe="36",binary="flag",moniker="34262",exit_status="0"
unc,Bdy,OneFunc,Bdy,OneFunc,Func,PlaneX,Func,PlaneX,Func,PlaneX,Func,PlaneZ,Func,PlaneZ,Func,PlaneY,Func,PlaneY,Mat,GasPiece,GasElement,PolyGas,GasMat,MatQ,MatQBarton,GasEOS,Gamma,Ir

heapstackpe0="104222,104222,104600,104600,104600,104600,104600,104600,104600,104600,104600,104600,104600,104600,104600,104600,104600,104600,104600,104600,104600",numpe="36",binary
utput/MemInfo,/Global/Mesh/Bdy,/Global/Mesh/Bdy/OneFunc,/Global/Mesh/Bdy,/Global/Mesh/Bdy/OneFunc,/Global/Mesh/Bdy,/Global/Mesh/Bdy/OneFunc,/Global/Mesh/Bdy,/Global/Mesh/Bdy/OneFunc,
nc,PlaneY,Func,PlaneY,Mat,GasPiece,GasElement,PolyGas,GasMat,MatQ,MatQBarton,GasEOS,Gamma,InitGas,GasPTRE,GasPTRE,Region,UniBdyReg,Region,OneFunc,Region,BoolReg,Region,BoolReg,Hydro,
```

Automatic Field Summary

5



#	tc_cpu_ParStructZoner	7
#	tc_cpu_ReshapeH	1
#	tc_cpu_ReshapeQ	1
#	tc_cpu_ReshapeU	1

Interesting Fields

a	binary	1
a	build_dt	2
a	class	2
a	classpath	2
a	cwd	1
#	date_mday	1
#	date_minute	6
a	date_month	1
#	date_second	7
a	date_wday	1
#	date_year	1
a	date_zone	1
a	deck_md5	4
#	exit_status	1
a	exit_sub	1
a	heapstackpe0	6
a	hostname	2
#	Hydro	7
a	index	1
a	info	1
a	insertTime	5
#	linecount	1
#	metricsVersion	1

heapstackpe0 ✕

6 Values, 100% of events

Selected Yes No

Reports

Top values Top values by time Rare values

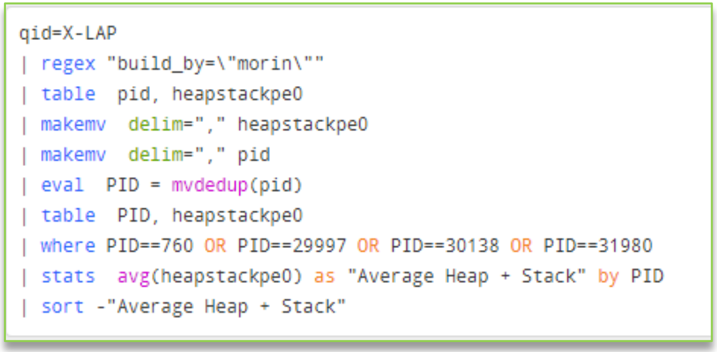
Events with this field

Values	Count	%
59745,59745,59808,59808,59808,59808,59808,59808,59808,59808,59808,59808,59808,59808,59808,59808,59808	2	28.571%
104221,104221,104599,104599,104599,104599,104599,104599,104599,104599,104599,104599,104599,104599,104599,104599,104599,104599,104599,104599,104599	1	14.286%
104222,104222,104600,104600,104600,104600,104600,104600,104600,104600,104600,104600,104600,104600,104600,104600,104600,104600,104600,104600,104600	1	14.286%
104420,104420,147536,147536,147536,147536,147536,147536,147536,147536,147536,147536,147536,147536,147536,147536,147536,147536,147536,147536,147536,147536,147536,147536,147536,147536,147536,147536,147536,147536,147536,147536,147536	1	14.286%
423216,423216,423216,423216,423216,423216,423216,423216,423216,423216,423216,423216,423216,423216,423216,423216,423216,423216,423216,423216,423216,423216,423216,423216,423216,423216,423216,423216,423216	1	14.286%
59744,59744,59807,59807,59807,59807,59807,59807,59807,59807,59807,59807,59807,59807,59807,59807,59807	1	14.286%

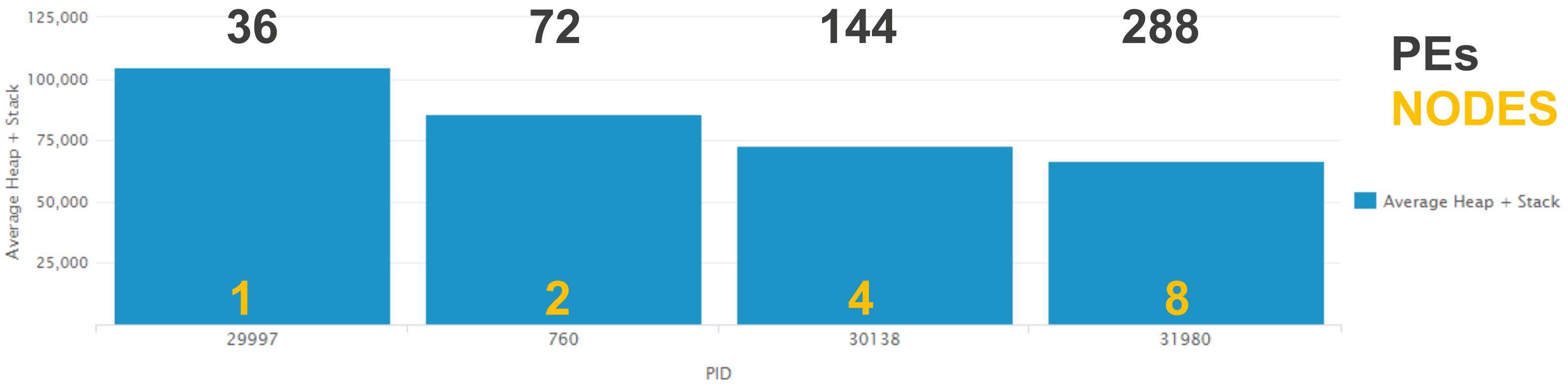
Splunk + MemInfo

Input: **sod3d_ndim64.flg**

Settings: **ppr:1:core**



```
qid=X-LAP
| regex "build_by=\"morin\""
| table  pid, heapstackpe0
| makemv  delim="," heapstackpe0
| makemv  delim="," pid
| eval  PID = mvdedup(pid)
| table  PID, heapstackpe0
| where PID==760 OR PID==29997 OR PID==30138 OR PID==31980
| stats  avg(heapstackpe0) as "Average Heap + Stack" by PID
| sort -"Average Heap + Stack"
```



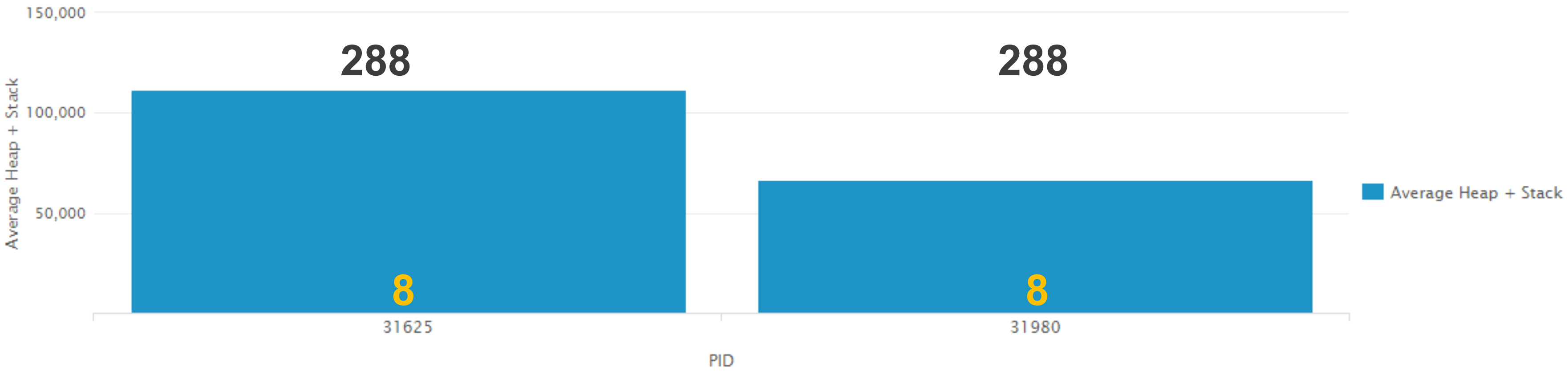
Splunk + MemInfo

sod3d_ndim128.flg

sod3d_ndim64.flg

PEs

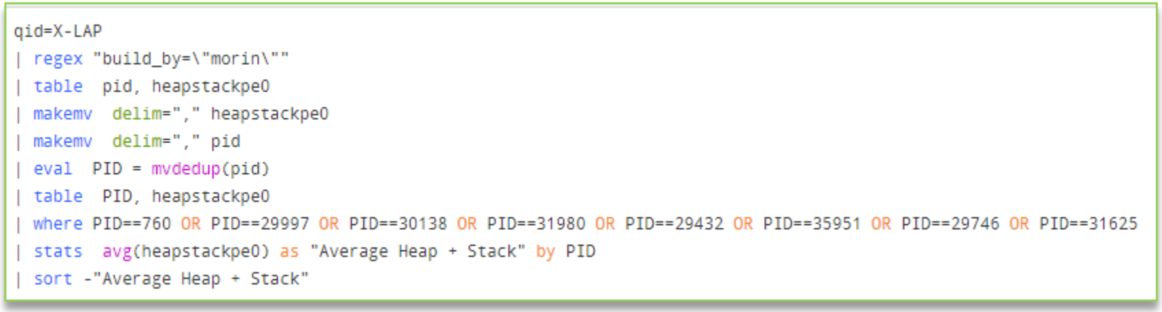
NODES



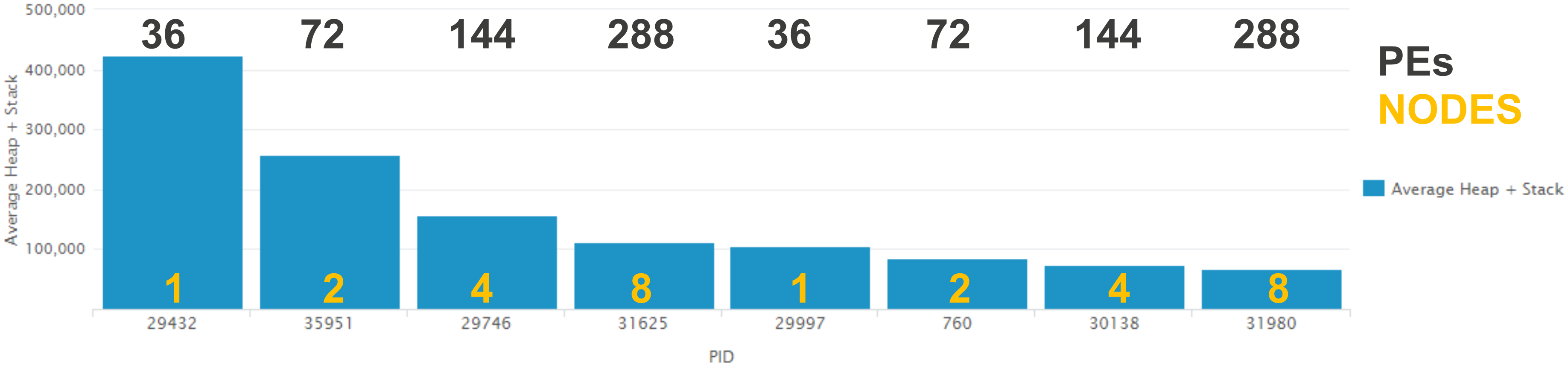
Splunk + MemInfo

sod3d_ndim128.flg

sod3d_ndim64.flg



```
qid=X-LAP
| regex "build_by=\"morin\""
| table  pid, heapstackpe0
| makemv  delim="," heapstackpe0
| makemv  delim="," pid
| eval  PID = mvdedup(pid)
| table  PID, heapstackpe0
| where PID==760 OR PID==29997 OR PID==30138 OR PID==31980 OR PID==29432 OR PID==35951 OR PID==29746 OR PID==31625
| stats  avg(heapstackpe0) as "Average Heap + Stack" by PID
| sort -"Average Heap + Stack"
```



Conclusions + Future Work

What do we have now?

- Toolbox:
  - Splunk = powerful
  - Potential to mine meaningful data
  - Workflow to get data from FLAG to Splunk

Where do we want to go?

  - Splunk?
  - Other data?

Data informed decisions:

Learn how to improve FLAG based on memory usage

Thank You XCP-1!

Thank you:

- Jimmy for hosting me
- Wendy + everyone I spoke with/listened to

Highlights:

- The lab for staying committed to students
- Designing Your Career Series (based on book)

Questions for you all:

- Have you worked anywhere besides LANL and how does it compare?
- What is one thing you wished you knew when prepping to finish undergrad?